

Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

RESUMO E OBJETIVOS

A Inferência Estatística é um tema da disciplina de MACS que permite otimizar o uso da tecnologia na criação de experiências estatísticas, em particular na obtenção de amostra e cálculo imediato de medidas estatísticas.

Nesta atividade pretende-se que os alunos, perante uma situação problema, executem simples estudo de inferência estatística em que são conhecidos os dados da população, embora os alunos não os observem. O parâmetro a estimar é o valor médio das alturas de 100 trabalhadores de uma empresa. Com recurso à aplicação Listas e Folha de Cálculo, da TI-Nspire CX, e numa página já predefinida, o aluno obterá uma amostra aleatória e introduzirá as fórmulas que lhe permitam calcular medidas estatísticas da sua amostra e, ainda, o cálculo do intervalo de confiança a 95%.

Em conclusão, com esta atividade pretende-se:

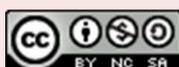
- Ver e rever os procedimentos num processo de Inferência Estatística.
- Recordar e reforçar os termos usado em Inferência Estatística: população, parâmetros, amostra, estatísticas, estimativa, intervalo de confiança.
- Consolidar os conceitos e interpretação de estimativa e intervalo de confiança.
- Estudar a importância da dimensão da amostra no processo de Inferência Estatística.
- Promover a discussão em sala de aula e a capacidade de análise.

MATERIAIS E PREPARAÇÃO

- TI-Nspire CX ou CX II-T
- Ficha do Aluno (uma por aluno)
- Ficheiro inferencia_alturas.tns

Esta tarefa deve ser realizada, preferencialmente, em grupo devendo cada grupo ter acesso a pelo menos uma unidade portátil (calculadora) TI-Nspire CX II ou ao respetivo software.

A tarefa pressupõe que os alunos já tenham conhecimento sobre Inferência Estatística, em particular, cálculo da estimativa do valor médio populacional, podendo, no entanto, também ser utilizada como introdução ao tema, devendo neste caso o professor reservar algum tempo inicial expor o processo de Inferência Estatística. Sugere-se que, num ou noutro caso, os alunos possam consultar o seu manual.



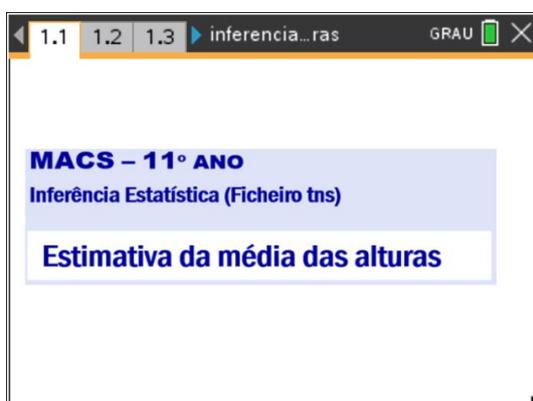
Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

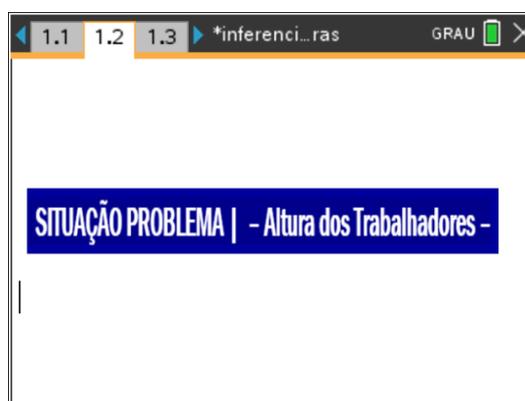
TAREFAS E INVESTIGAÇÕES PARA OS ALUNOS

Considera-se que o contexto dado a esta tarefa, situação problema, merece uma leitura cuidada por parte dos alunos e uma discussão em grande grupo turma, em particular quanto às perceções acerca da importância da dimensão da amostra e do que poderá ser considerada uma boa estimativa. Poderá ser levantada aqui a questão, como saber se a estimativa é boa se não sabemos o parâmetro a estimar! É aqui que a Matemática é o garante da “ciência”!

Colocam-se de seguida várias capturas de ecrã das páginas do documento **inferencia_alturas.tns**:



Página 1.1 - Tema



Página 1.2 - Problema “Altura dos Trabalhadores”



Página 1.3 - Simulador



Página 1.3 - Simulador vista computador

QUESTÃO 1.

Nesta questão pretende-se que os alunos, lançados no desafio de encontrar um processo aleatório, sintam dificuldades e discutam entre eles o que será verdadeiramente um processo aleatório de seleção de uma amostra. O professor poderá levantar aqui a questão das tabelas de números

Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

aleatórios e o facto de num processo computacional existir um algoritmo e uma semente, e por isso ser habitual designar números pseudoaleatórios. Poder-se-á aproveitar, aqui, para referir a necessidade de cada aluno ter de atribuir uma semente diferente ao algoritmo gerador de números aleatórios da sua TI-Nspire CX, executando numa página de calculadora a instrução *randseed* seguida de um qualquer número.

Nesta primeira experiência dos alunos, não deve constituir preocupação o facto de a amostra ser verdadeiramente aleatória, dado o objetivo ser a construção do simulador e a análise dos resultados. No entanto, essa referência poderá ser feita pelo professor se considerar oportuno.

No ecrã ao lado encontra-se um exemplo de uma possível situação criada por um aluno. Não esquecer que após

inserir a ordem dos elementos da população, ou após qualquer alteração, terá que se clicar **ctrl** + **R**.

(Nota: Para obter os valores das medidas estatística na forma decimal pode-se multiplicar, em cada célula, a fórmula por 1.0, assim o resultado surgirá na forma decimal.)

A ord...	B amostra1	C
=	=amostra	
1	21	160
2	45	167
3	67	160
4	12	168
5	98	159

Dimensão Amostra	10
Média Amostral	166.9
DP Amostral	

C4 =mean(amostra1)· 1.

QUESTÃO 2.

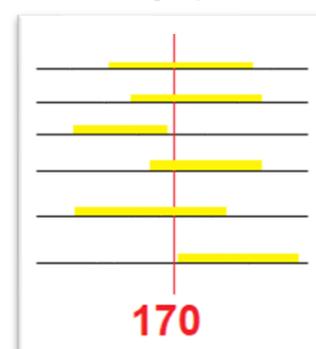
Nesta questão os alunos, usando a fórmula para o intervalo de confiança a 95% para o valor do médio populacional, a partir da média e do desvio-padrão amostral, inserem as fórmulas que permitem calcular os extremos do intervalo, e de seguida a fórmula para obterem a amplitude do intervalo. Os alunos terão que usar a referência relativa a outras células nas fórmulas dos extremos do intervalo de confiança, conforme é possível observar no ecrã ao lado.

Após o registo dos intervalos obtidos, os alunos, em grupo de trabalho e/ou em grupo turma, deverão analisar os vários intervalos obtidos na turma, realçando-se aqui a amplitude e a variação dos limites inferiores e superiores dos intervalos de confiança. Poderá ser pertinente representarem-se os vários intervalos na reta real, dispondo-se várias retas horizontais paralelas e centradas num mesmo valor. Esta representação permitirá uma melhor perceção da variabilidade dos intervalos, e servirá ainda para uma análise posterior.

D
10
Média Amostral
166.9
DP Amostral

Limite Inferior=	161.33
Limite Superior=	172.47

D3 =c4 - $\frac{1.96 \cdot c6}{\sqrt{c2}}$



Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

QUESTÃO 3.

Nesta questão os alunos deverão calcular os parâmetros valor médio e desvio padrão populacional. Para tal, na mesma folha de cálculo, devem inserir nas células respetivas as funções da TI-Nspire CX que permitem obter os valores, a saber:

- célula E3 =**mean**(alturas)
- célula E5 =**stdevpop**(alturas)

De notar que para obter os valores na forma decimal pode-se, por exemplo, multiplicar cada fórmula por 1.0 .

Após o cálculo do parâmetro os alunos deverão comparar com a estimativa calculada na sua amostra e verificar se o

intervalo de confiança contém o parâmetro. Note-se que, embora possa ser equivalente, na verdade ao afirmamos que é o intervalo de confiança que contém o valor médio populacional, e não o contrário, estamos a reforçar a importante ideia que quem varia é o intervalo de confiança, e não o parâmetro, embora desconhecido.

A resposta à questão “Qual a percentagem IC que contém o parâmetro?”, permite introduzir o conceito de nível de confiança e a sua compreensão. Espera que com a última questão os alunos apontem estratégias como aumentar a dimensão da amostra, sendo esta uma oportunidade para o professor introduzir outros níveis de confiança.

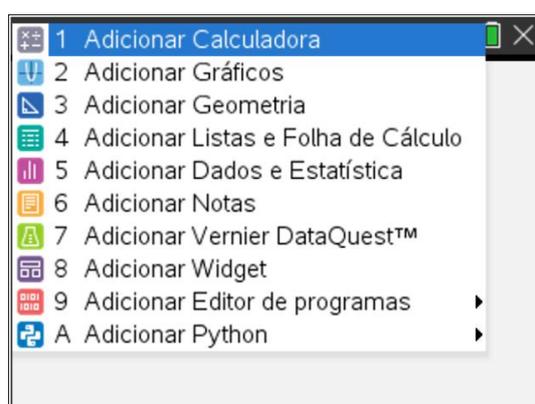
	Intervalo Confiança	POPULAÇÃO
1	Limite Inferior=	Valor Médio (μ)
2		
3	161.33	167.12
4	Limite Superior=	DP Populacional (σ)...
5	172.47	8.95464

E3 =mean(alturas)· 1.

QUESTÃO 4.

Nesta questão, fazendo uso do simulador de estimativas acabado de criar, vão gerar amostras pseudoaleatórias de várias dimensões, 30, 40 e 50, e registar para cada simulação a sua dimensão, a estimativa, o intervalo de confiança e a sua amplitude.

Abaixo colocam-se ecrãs que exemplificam o processo acima referido:



Adicionar Página Calculadora (**ctrl** + **doc**)



Indicar semente do algoritmo aleatório

Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

GERAR AMOSTRAS DE DIMENSÃO 30

A	ord...	B amostra1	C
		=	= amostra
1	21	160	Dimensão Amostra
2	45	167	10
3	67	160	Média Amostral
4	12	168	166.9
5	98	159	DP Amostral

A ordem:=randint(1,100,30)

Gerar 30 números inteiros aleatórios

Perda de dados

Esta operação substituirá os dados das colunas actuais. Deseja continuar?

OK Cancelar

Substituir os dados anteriores pelos novos

A	ord...	B amostra1	C
		=	= amostra
1	48	178	Dimensão Amostra
2	62	168	30
3	12	168	Média Amostral
4	92	165	165.067
5	51	157	DP Amostral

C2 =dim(amostra1)

ATENÇÃO: Clicar **ctrl** + **R** para atualizar

	E
1	Intervalo Confiança POPULAÇÃO
2	Limite Inferior= Valor Médio (μ)
3	162.442 167.12
4	Limite Superior= DP Populacional (σ)...
5	167.691 8.95464

E3 =mean(alturas) * 1.

Surgirá nova amostra e medidas recalculadas

GERAR AMOSTRAS DE DIMENSÃO 40

A	ord...	B amostra1	C
		=	= amostra
1	80	159	Dimensão Amostra
2	90	174	40
3	10	158	Média Amostral
4	30	158	166.275
5	15	162	DP Amostral

A1 =80

Nova amostra de dimensão 40 gerada

	E
1	Intervalo Confiança POPULAÇÃO
2	Limite Inferior= Valor Médio (μ)
3	163.801 167.12
4	Limite Superior= DP Populacional (σ)...
5	168.749 8.95464

E3 =mean(alturas) * 1.

Medidas estatísticas recalculadas

Estimativa da média das alturas

Eduardo Cunha
Raul Aparício Gonçalves

GERAR AMOSTRAS DE DIMENSÃO 50

A ord...	B amostra1	C
1	28	178 Dimensão Amostra
2	89	185 50
3	62	168 Média Amostral
4	70	167 168.44
5	58	178 DP Amostral

Nova amostra de dimensão 50 gerada

3	165.991	167.12
4	Limite Superior=	DP Populacional (σ)...
5	170.889	8.95464
6	Amplitude IC	
7	4.89884	

Medidas estatísticas recalculadas

Após os registos dos vários intervalos de confiança por parte de todos os alunos da turma, o professor poderá promover a análise dos mesmos, representando os vários intervalos obtidos por dimensão da amostra num esquema semelhante ao apresentado na figura ao lado.

Nesta situação, sendo já conhecido o parâmetro a estimar, o valor médio das alturas da população, os intervalos podem ser representados em retas reais paralelas centradas em μ . Desta forma, poder-se-á observar o número de intervalos de confiança que contêm o parâmetro μ e as suas amplitudes, em função da dimensão das amostras.

Este esquema, análogo ao da figura ao lado, permitirá refletir sobre a influência da dimensão da amostra no nível de confiança, percentagem de intervalos de confiança que contêm o parâmetro a estimar, e na amplitude do mesmo intervalo (que se relaciona com a margem de erro). Para que este estudo faça sentido o professor deve ter o cuidado de recolher um número significativo de amostras para cada dimensão, sugere-se no mínimo 100 para cada caso.

